Beyond Assessment:
Accelerating Simulation's Readiness Benefits with Competency Profiles

Benjamin Bell, Matthew Ross, Debbie Brown, Fritz Ray
Eduworks Corporation
Corvallis, OR
[benjamin.bell|matthew.ross|debbie.brown|fritz.ray]@eduworks.com

ABSTRACT

Performance data from simulations can provide readiness insights, but interpreting what simulation outputs tell us about an operator's readiness remains a challenge. Programs such as the USAF's Pilot Training Transformation (PTT) make extensive use of simulation, and PTT is exploring competency-enabled performance tracking that can map simulator assessments to specific airmanship competencies. Eduworks is digitizing PTT competency maps (characterizing the proficiencies required by pilots for several stages of the training pipeline) into standardized, machine-readable competency frameworks. These frameworks can help identify what a pilot's current capabilities are, isolate any gaps, inform decisions about advancement and track selection, accumulate training effectiveness data, and provide a performance baseline to assist pilots, instructors, and squadrons through the training pipeline and into operational assignments. This work incorporates the open-source Competency and Skills System (CaSS), developed by Eduworks for DOD, to capture competency maps from multiple undergraduate flying training programs as digital frameworks, associate assessment data from multiple sources (instructor grading sheets, automated metrics, computer-based training, examinations) with corresponding competencies, and define how evidence of one competency contributes to calculations of any related competencies. For visualization, we are creating "competency profiles" – on-demand, evidence-based assessments of a pilot's competencies at a particular point in time. In this paper, we illustrate the process of creating digitized frameworks and associating them with performance data, and present current examples of how CaSS is helping accelerate adoption of competency-enabled approaches to pilot training. We conclude with a discussion of how these techniques apply more generally to readiness in the military and beyond.

ABOUT THE AUTHORS

Dr. Benjamin Bell is the COO of Eduworks, where he leads AI programs for simulation, training, and decision support. His research has addressed the use of simulation for training across a spectrum of applications, including K-12, higher education, military, and national security training. He has held faculty positions, chief executive positions in industry, and leadership roles for multiple conferences. He holds a PhD from Northwestern, and a BSE from the University of Pennsylvania.

Matthew Ross is Director of Government Solutions at Eduworks, where he leads AI implementation throughout the DOD and other USG agencies. He was previously a U.S. Air Force F-15E Instructor Pilot, where he steered the technical and programmatic development of multiple projects related to competency-based learning, performance tracking, and using analytics to predict career success. Matthew has an MBA from Duke, an MA from the Naval Post Graduate School, and a BS from USAFA.

Debbie Brown is the Director of Accounts at Eduworks and since joining Eduworks has acted as Learning Engineer, Instructional Designer, Technical Lead, and Project Manager across numerous projects. She has been active in technology-enabled learning research and development for 25 years in academic, workforce, and government/military training applications with a focus on the empowerment of non-technical users through user-centered design, standards-based models and interoperability, and semi-automated AI-assisted workflows. She holds a M.S. in Instructional Technology and Design, and a B.S. in Computer Engineering, both from Mississippi State University.

Fritz Ray is Eduworks' CTO and has been the solution architect leading Eduworks' CaSS program since inception. He oversees competency framework development, development of tools for data ingestion, architecture, and integration of CaSS. Mr. Ray has over fifteen years experience architecting, designing, and leading development of software used by our US Government clients and industry partners in the fields of readiness, training, and talent management. He holds a B.S. in Software Engineering Technology from the Oregon Institute of Technology.

Beyond Assessment: Accelerating Simulation's Readiness Benefits with Competency Profiles

Benjamin Bell, Matthew Ross, Debbie Brown, Fritz Ray
Eduworks Corporation
Corvallis, OR
[benjamin.bell | matthew.ross | debbie.brown | fritz.ray]@eduworks.com

INTRODUCTION

Training simulators have been in widespread use across military organizations for decades, and both their adoption and their sophistication continue to grow. The performance data that these devices generate can provide a wealth of readiness insights, but interpreting what simulation outputs tell us about an operator's readiness remains a challenge. In recent work, we have been working with the U.S. Air Force Pilot Training Transformation (PTT) program, which integrates traditional flying sorties with VR-enabled ground-based training devices and data-driven proficiency tracking to achieve training efficiencies, improve readiness, and increase throughput. PTT, and comparable programs in the U.S. Navy and U.S. Army, generate extensive simulator output data. To help make sense of this data, PTT is exploring competency-enabled performance tracking that can map simulator assessments to specific airmanship competencies.

In this paper, we outline our process for capturing the proficiencies required by pilots as digital competency maps, providing standardized, machine-readable competency frameworks. We discuss our approach for associating the competencies within these digitized frameworks with simulator output data and illustrate how we calculate on-demand "competency profiles" providing evidence-based assessments of a pilot's current capabilities. We conclude with a discussion of how these techniques apply more generally to readiness in the military and beyond.

SIMULATION AND THE PERFORMANCE DATA PROBLEM

The Air Force employs an extensive inventory of flight training devices across a spectrum of fidelities, training communities, and platforms. Every device, from desktop part-task trainers to full mission simulators to airplanes, generates volumes of data. The challenges of manually mapping raw simulator data to corresponding competencies, even with derived metrics from tools such as the AF's Performance Evaluation Tracking System (PETS), have to-date limited the ability to capture the proficiencies being demonstrated throughout a training event. Broad enhancements to training efficacy are achievable given a more comprehensive and persistent learner profile and comprehensive picture of the competencies being trained (Goldberg, *et al.*, 2021; Selmanagić & Simbeck, 2022; Walcutt, Harley, Spohn, & Bockelman, 2022). Below we introduce an approach to achieving this kind of human performance data management that offers a foundation to enhance a broad spectrum of simulation-based training.

OBJECTIVES OF THE CURRENT WORK

The goal of this work is to provide Air Force training stakeholders with competency insights to enhance training. Our approach is to develop and integrate competency frameworks in the Competency and Skills System (CaSS) and map data coming from PTT to track individual Airman competency profiles. CaSS (Robson, et al., 2021) is a mature, TRL 8 product developed by Eduworks since 2015 with DOD investment. It provides a competency engine that ingests learning frameworks and connects those frameworks to individual learners. These frameworks can help identify what a pilot's current capabilities are, isolate any gaps, inform decisions about advancement and track selection, accumulate training effectiveness data, and provide a performance baseline to assist pilots, instructors, and squadrons through the training pipeline and into operational assignments.

MAPPING SCENARIOS AND ASSESSMENTS TO COMPETENCIES

CHECK-SIX demonstrates the application of performance indicators to calculating a common proficiency profile. An important step in this process is to establish a digital relationship among scenarios, maneuvers, and the corresponding competencies. In the initial phase of this work, we began this mapping step with a manual process to extract structured competencies from existing SUPT frameworks, cross-walk with the learning objectives that accompany each PTT scenario as defined in the T-6 syllabus, and identify the pool of potential common competencies for the scenario events (for instance, the competency "Perform Enroute Climb" would be associated with learning objectives "Altitude control", "Airspeed control", and "Basic Aircraft control and performance (Climbs, turns, and descents)". These mappings are defined in CaSS as node relations that can be traversed during CaSS profile computation. We are now exploring a more sophisticated approach to automatically identify skills indicated by the telemetric scenario data.

Similarly, the maneuvers identified in instructor gradesheets are extracted and stored in CaSS, cross-walked with the tasks in the SUPT framework, and aligned to common competencies that support maneuver-based proficiency computations. Figure 1 illustrates these mappings, depicting scenario outputs and instructor grade sheets associated with individual competencies and sub-competencies in a CaSS framework.

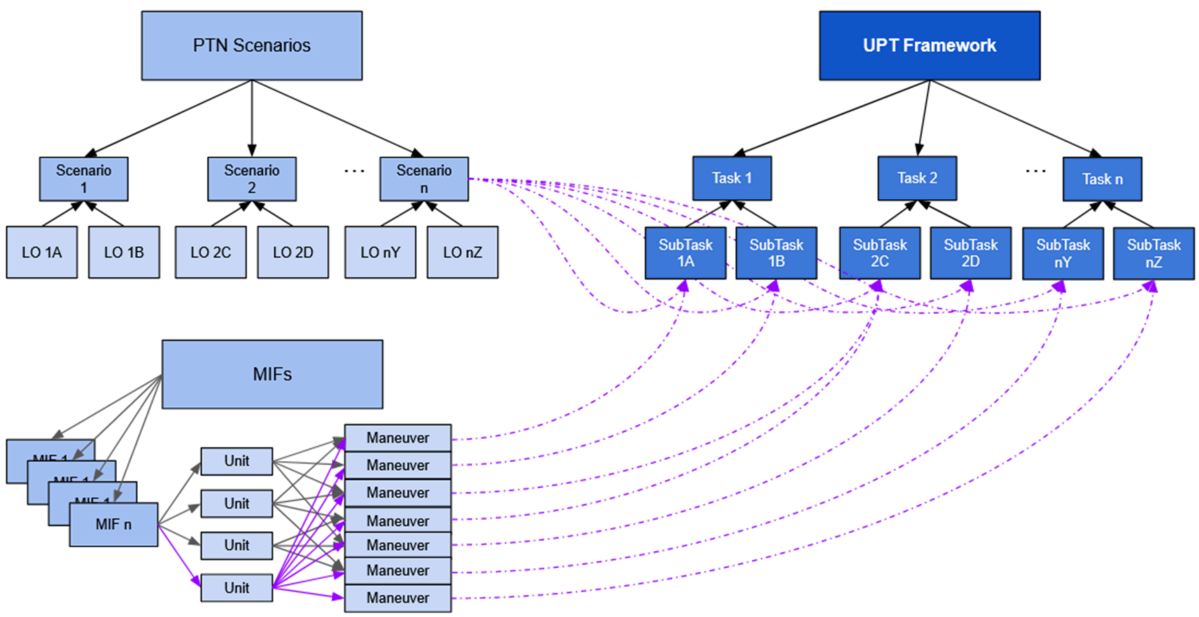

Figure 1. Mapping inputs to CaSS competencies: scenario outputs (top left); grade sheets (bottom left)

CHECK-SIX ARCHITECTURE

As depicted in Figure 2, the principal components of the architecture are:

- Competency and Skills System (CaSS)
- CHECK-SIX Assertion Generator
- CHECK-SIX Assertion Processor
- CHECK-SIX Application Programming Interface (API)
- CHECK-SIX Dashboard to exercise and demonstrate the system

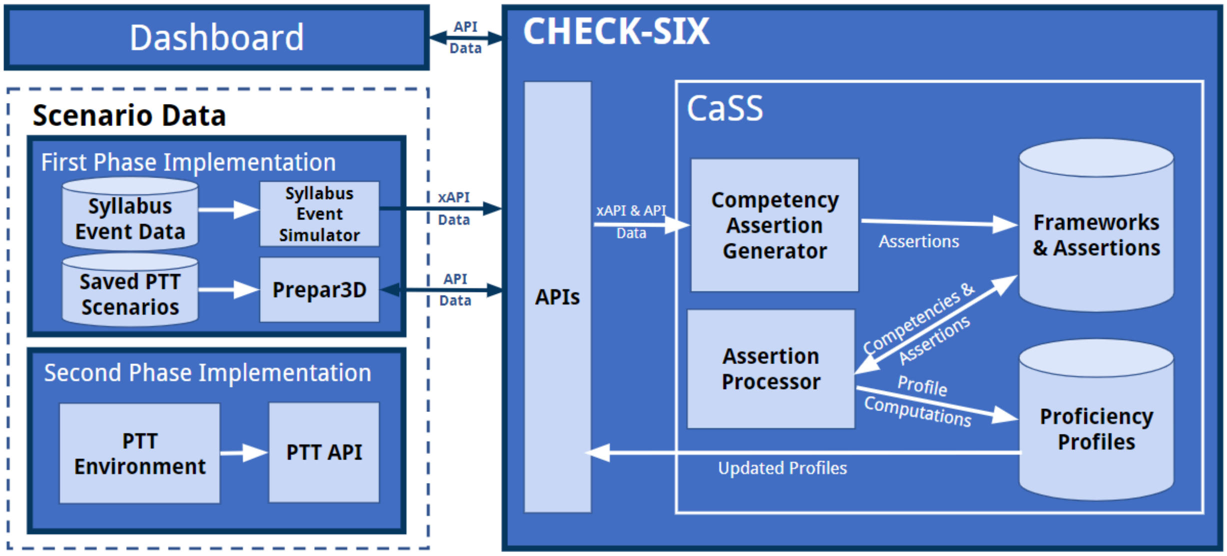

Figure 2. Schematic Diagram of the CHECK-SIX Architecture.

CaSS is a key component of CHECK-SIX that processes data generated from PTT and maps those data into competency assertions. CaSS includes three primary capabilities. The first is used to author, manage, and share competency frameworks in both human-readable and machine-actionable formats. The second collects and stores evidence-based assertions about the competencies that have been demonstrated by a pilot flying a PTT scenario. These assertions can derive from multiple sources: instructor grading sheets, automated metrics, computer-based training, examinations, and Experience API (xAPI) activity reports. CaSS frameworks then govern how evidence of one competency contributes to calculations of any related competencies. These assertions are stored in a uniform format that enables them to be used in knowledge tracing algorithms, skill gap analysis, pathway recommenders, and skills-based advancement and tracking. The third capability of CaSS enables custom algorithms to be applied to compute Airman profiles that can then be visualized within a CaSS interface and accessed by other systems. This enables algorithms and AI-based models to be implemented and tested prior to being incorporated into live systems. CaSS is open-source and licensed under an Apache 2.0 license.

The **Assertion Generator** (AG) provides the services required to process pilot performance and experience inputs submitted via the CHECK-SIX API. The AG is responsible for handling and transforming incoming data, deciphering which PTT component event is being reported for which Airman, and deciding how and what competency assertions need to be made for the Airman's performance record. CHECK-SIX is initially scoped to ingest PTT scenario outputs as well as instructor pilot assessments structured around Maneuver Item File (MIF) items and submitted via the CHECK-SIX API. The PTT scenario outputs provide raw instrument data feeds and are captured as evidence of exercise experience on the relevant competencies using CaSS assertions. The MIF-structured instructor pilot assessments (grade sheets) are captured as performance evaluations on the relevant competencies using CaSS assertions as well. In addition, the native CaSS xAPI Adaptor monitors an xAPI activity feed when there is a Learning Record Store (LRS) configured. CaSS generates assertions based on xAPI-reported assessment outcomes that are associated with competencies. All three types of evidence are evaluated and incorporated into the Airman's competency-based computed profile via CaSS and accessed via the CHECK-SIX API.

Part of the power of CaSS is that, once a competency framework is properly defined, the system computationally manages the ripple effects that one competency assertion exerts on all other competencies linked via a defined relation (subtask, equivalence, narrows, expands, requires). The competency framework created for CHECK-SIX thus represent a significant improvement over existing competency approaches by capturing the relationships among the elements composing the top-level competency, tying each competency to the evidentiary data source(s) that document its accomplishment and by relating activity in a training-specific framework to a common competencies representation.

Assertion Processing is the process by which all activity relative to a specific competency in a pilot's competency profile is reviewed and evaluated in an automated estimation of mastery. This processing is natively handled within CaSS as part of the learning profile feature. The Assertion Processor computes an Airman's current state for each of the competencies exercised within a scenario session based on the assertions that have been generated for that Airman's training event. The logic for processing assertions is based on historical record (as assertions) and known definitions and relationships (as competency frameworks), but it is a computation of state for a particular point in time and is not static. It is the computational analysis of assertions that provides the proficiency snapshots for an Airman, and it is the continuous processing of assertions that maintains the CHECK-SIX dashboard updated in real time as a scenario is performed.

The CHECK-SIX **API** supports the communication of PTT performance information to and from CaSS, which enables a centralized view of an Airman's competencies in real time. The representation of an Airman's currently demonstrated state of training is computed in real time, relying on CaSS's understanding of the overall training requirements, the overall skills, knowledge, and abilities expected of the Airman, the tasks and skills that are executed in particular exercises, and the collection of exercise performance history and associated evaluations from instructors.

The API provides the communication layer among internal components and serves as a point of integration with external systems. All messages are implemented as a representational state transfer (RESTful) interface in which the requests include information required to complete the requested operation and the responses include the requested data. Requests and responses include headers and/or parameters that establish the type of HTTP connection, metadata, authorization, caching, and status codes as appropriate. All payloads are represented as Javascript Object Notation (JSON). The API also includes messages that manage the CHECK-SIX data processing pipeline and handles requests to support the dashboard for monitoring or observing learner profiles. The CaSS API is incorporated into CHECK-SIX and is used for lower-level operations on frameworks, competencies, assertions, and profiles.

The CHECK-SIX **Dashboard** displays changes in an Airman's competency-based profile as syllabus events are completed and as evidence is ingested from avionics computer-based instruction and from instructor evaluations via digital gradesheets; the Dashboard also displays performance histories. The Dashboard is designed to accommodate additional inputs via the PTT API as they become available, as well as xAPI activity statements. We also incorporated a training event simulator and animated illustration of data movement into the dashboard layout. The training event simulator issues xAPI activity statements as a surrogate until realtime xAPI data is published. The animated illustration depicts xAPI statements being sent to the LRS and being detected and asserted by CaSS. Finally, CaSS profile updates are dynamically rendered in the competency view for that Airman.

VISUALIZING COMPETENCY PROFILES

As depicted in Figure 3, the CHECK-SIX Dashboard serves as an interface that can demonstrate CHECK-SIX architectural component functionality. The left-hand side of the dashboard serves as a test harness to simulate xAPI data coming from training systems within the PTT ecosystem. This feature was added to the interface after PTT adopted the xAPI standard to share performance data among systems. We defined xAPI activity identifiers to correspond to the tasks, knowledge and skills to be exercised across PTT syllabus events. The test harness generates a number of competency-based xAPI performance reports for each syllabus event. In our simulated PTT ecosystem, an LRS is used to collect the disparate xAPI statements, and CaSS is configured to query the xAPI activity feed and generate assertions when it detects that there has been a change in mastery status of a competency for a particular Airman. As assertions are generated, CaSS recomputes the Airman's competency profile to estimate mastery state for each competency. The dashboard on the right side presents dynamic changes in the competency profile. This simple view was constructed as a technology demonstration, but we discuss below other types of visualizations that this approach enables.

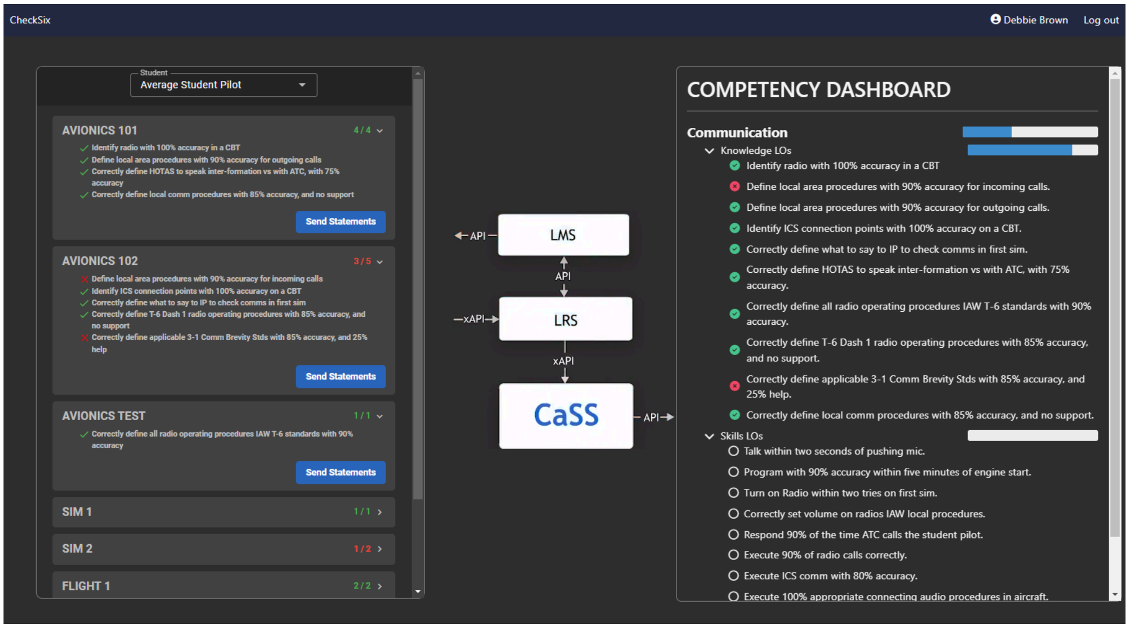

Figure 3. Dashboard showing training events (left), data flow (center), and competency profile (right).

INTEGRATION

Integrating CHECK-SIX with PTT requires several steps in an evidentiary pathway that begins with the simulation environment, governed by models and data from the simulation event stream, and continues through the assertion of competence for one or more competencies. This information pipeline requires the integration of several components, each transforming inputs to outputs governed by one or more standards, as summarized in Table 1. The result of this process is the integration of these elements into the PTT device shown in Figure 4.

Table 1. Elements of the CHECK-SIX Integration Pipeline

Element	Inputs	Outputs	Standards
Simulation	Sensors and Models of Tasks, Maneuvers, Task Performance, Outcomes	Baseline performance data	DIS, HLA
Assessment Engine	Models of Events, Assessments, Measures, Thresholds, and Sequences	Assessed performance data	QTI
LMS	Scenarios, Lessons	xAPI Statements	SCORM, CMI5, xAPI
LRS	Activities	xAPI history of performance	xAPI
CaSS	LO, Tasks, Competencies, Outcomes, other models of capability/performance	Assertions of capabilities	IEEE 1484.20.3, CTDL, schema.org
Profile Calculator	Models for aggregating assertions into conclusions about individual capabilities	Conclusions used to provide decision support, adaptive learning, or learner diagnostics	
Dashboard	Identifying relevant analytics and insights	Visualizations depicting information flow/profile updates	

Figure 4. Portable Pilot Transformation Training station hosting CHECK-SIX.

PERFORMANCE INSIGHTS AND PILOT TRAINING BENEFITS

The goal of CHECK-SIX is to objectively determine student performance at individual and group levels. With granular data on student performance on every attempt of learning every piece of required knowledge, and performing every skill, task and learning objective, CaSS builds an in-depth performance tracking system previously unavailable in pilot training. Specifically, CaSS accomplishes root cause analysis on a student's failures automatically and across the student's entire time in training. For instance, Figure 5 below demonstrates a theoretical student pilot struggling in the core competency of "Mission Execution". Through CHECK-SIX curriculum design and tracking, an instructor can readily discover that the student is struggling with the skill "Application of Procedures", the task "Perform Emergency Procedures", the sub-task "Manage tasks IAW MATL" and finally the learning objective "Recognize Indications of non-critical emergencies." This is not a failure in knowledge and execution in one sortie, but across every attempt in the entire syllabus, trend data that would be impossible for even the most inquisitive instructor to discover about a student. This root cause analysis allows the instructor to give individualized training to each student pilot on the pilot's weakest areas - resulting in more effective training.

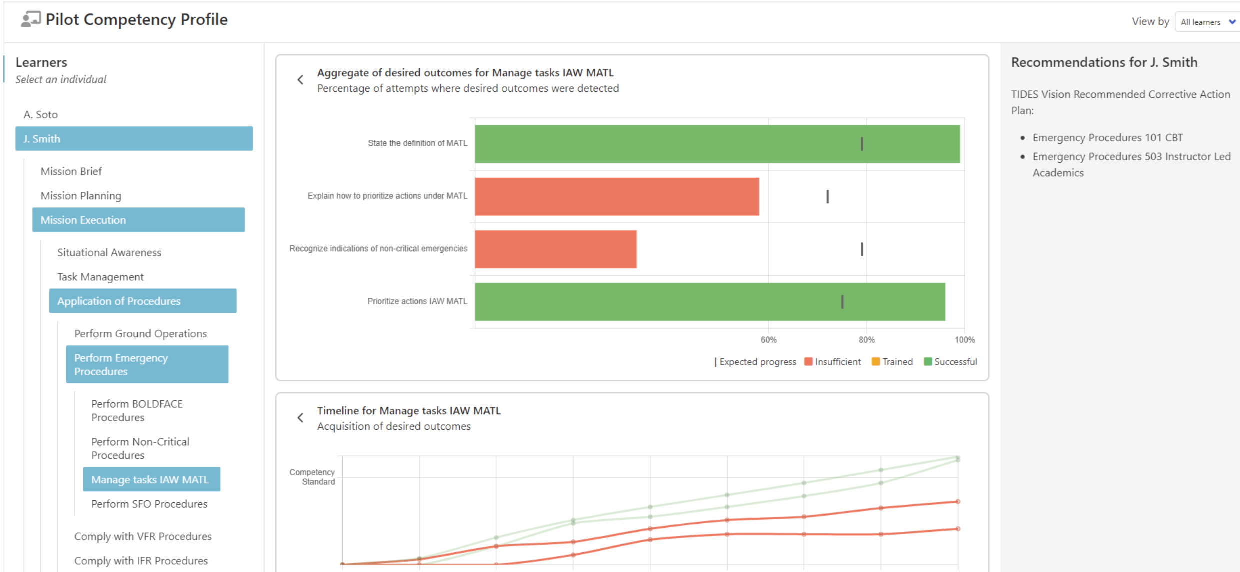

Figure 5. Current work: Training Effectiveness dashboard driven by CHECK-SIX outputs.

CaSS likewise combines all of this detailed data for an entire group (whether a class of 20 student pilots or all 1500 student pilots that graduate from UPT each year). This group data in near-real time feeds objective performance metrics back to 19AF leadership to locate shortcomings or overtraining in the syllabus or content development, and training developers can then promptly address those issues. This group view of objective performance at the detailed level allows 19AF to create more efficient and effective syllabi and content.

CONCLUSION AND FUTURE WORK

Simulator data can provide readiness insights, but requires more sophisticated means of interpreting the data than are currently available. In this paper we reported current work in digitizing pilot competency maps into standardized, machine-readable competency frameworks that can help identify what a pilot's current capabilities are. We described the open-source Competency and Skills System (CaSS) as a tool for representing competency maps as digital frameworks and for associating assessment data from multiple sources (instructor grading sheets, automated metrics, computer-based training, examinations) with corresponding competencies. We presented a current example, called CHECK-SIX, of how CaSS is helping accelerate adoption of competency-enabled approaches to pilot training.

This work is continuing and we anticipate the program will yield evidence of training outcome improvements. Beyond UPT, by expanding CHECK-SIX to the formal training units, the Air Force could gain the same detailed performance insights about its individual pilots and syllabi. By further expanding CHECK-SIX to operational units, the Air Force would gain both objective, outcome-based, readiness assessments at a more detailed level, and objective success and failure metrics to update ascension and aircraft assignments. Further, there is no reason to limit these advancement to pilots, who represent just 4% of total Air Force active duty professionals, when every other technical profession (maintenance, air traffic control, special operations) would also dramatically benefit from the same program. Finally, this approach is readily generalizable to other Services as well as any non-military organization that collects training data and requires on-going visibility into workforce readiness.

ACKNOWLEDGEMENTS

This work was supported by the US Air Force under contract FA8650-22-C-6025. The authors wish to acknowledge the guidance and support of the AF sponsor, Dr. Wink Bennett. The views expressed are those of the authors and do not reflect the official policy or position of the US Air Force, Department of Defense or the US Government.

REFERENCES

Department of the Air Force (2022). Competency Modeling. Air Force Handbook 36-2647, February 8, 2022.

Goldberg, B., Owens, K., Gupton, K., Hellman, K., Robson, R., Blake-Plock, S., & Hoffman, M. (2021). Forging competency and proficiency through the synthetic training environment with an experiential learning for readiness strategy. In Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

Robson, E., Robson, R., Buskirk, T., Ray, F., & Owens, K. P. (2021). An Experiential Competency Application Framework. In International Conference on Human- Computer Interaction (pp. 121-134). Cham: Springer.

Selmanagić, A., & Simbeck, K. (2022). Designing Granular Competency Frameworks for Adaptive Learning on the Example of Naïve Bayes Classifiers. In Proceedings of DELFI Workshops 2022 (p. 137).

Walcutt, JJ, Harley, T., Spohn, J., & Bockelman, P. (2022). Modernizing High-End Flight Training for the Contested Fight. In Proc. of the 2022 Interservice/Industry Training Simulation and Education Conference (I/ITSEC).